

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Wahl, Geoffrey M
O'Gorman, Stephen V
- (ii) TITLE OF INVENTION: FLP-MEDIATED GENE MODIFICATION IN
MAMMALIAN CELLS, AND COMPOSITIONS AND CELLS USEFUL
THEREFOR
- (iii) NUMBER OF SEQUENCES: 4
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
 - (B) STREET: 444 South Flower Street, Suite 2000
 - (C) CITY: Los Angeles
 - (D) STATE: California
 - (E) COUNTRY: USA
 - (F) ZIP: 90071
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 08/484,324
 - (B) FILING DATE: 07-JUN-1995
 - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Reiter, Stephen E
 - (B) REGISTRATION NUMBER: 31,192
 - (C) REFERENCE/DOCKET NUMBER: P41 9984
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: (619) 546-4737
 - (B) TELEFAX: (619) 546-9392

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1380 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (vii) IMMEDIATE SOURCE:
 - (B) CLONE: NATIVE FLP
- (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) LOCATION: 1..1269
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATG CCA CAA TTT GAT ATA TTA TGT AAA ACA CCA CCT AAG GTG CTT GTT
 Met Pro Gln Phe Asp Ile Leu Cys Lys Thr Pro Pro Lys Val Leu Val
 1 5 10 15

CGT Arg	CAG Gln	TTT Phe	GTG Val 20	GAA Glu	AGG Arg	TTT Phe	GAA Glu	AGA Arg 25	CCT Pro	TCA Ser	GGT Gly	GAG Glu	AAA Lys 30	ATA Ile	GCA Ala	96
TTA Leu	TGT Cys	GCT Ala 35	GCT Ala	GAA Glu	CTA Leu	ACC Thr	TAT Tyr 40	TTA Leu	TGT Cys	TGG Trp	ATG Met	ATT Ile 45	ACA Thr	CAT His	AAC Asn	144
GGA Gly	ACA Thr 50	GCA Ala	ATC Ile	AAG Lys	AGA Arg	GCC Ala 55	ACA Thr	TTC Phe	ATG Met	AGC Ser	TAT Tyr 60	AAT Asn	ACT Thr	ATC Ile	ATA Ile	192
AGC Ser 65	AAT Asn	TCG Ser	CTG Leu	AGT Ser	TTC Phe 70	GAT Asp	ATT Ile	GTC Val	AAT Asn	AAA Lys 75	TCA Ser	CTC Leu	CAG Gln	TTT Phe	AAA Lys 80	240
TAC Tyr	AAG Lys	ACG Thr	CAA Gln	AAA Lys 85	GCA Ala	ACA Thr	ATT Ile	CTG Leu	GAA Glu 90	GCC Ala	TCA Ser	TTA Leu	AAG Lys	AAA Lys 95	TTG Leu	288
ATT Ile	CCT Pro	GCT Ala 100	TGG Trp	GAA Glu	TTT Phe	ACA Thr	ATT Ile 105	ATT Ile 105	CCT Pro	TAC Tyr	TAT Tyr	GGA Gly 110	CAA Gln	AAA Lys	CAT His	336
CAA Gln	TCT Ser	GAT Asp 115	ATC Ile	ACT Thr	GAT Asp	ATT Ile 120	GTA Val	AGT Ser	AGT Ser	TTG Leu	CAA Gln 125	TTA Leu	CAG Gln	TTC Phe	GAA Glu	384
TCA Ser 130	TCG Ser	GAA Glu	GAA Glu	GCA Ala	GAT Asp	AAG Lys 135	GGA Gly	AAT Asn	AGC Ser	CAC His	AGT Ser 140	AAA Lys	AAA Lys	ATG Met	CTT Leu	432
AAA Lys 145	GCA Ala	CTT Leu	CTA Leu	AGT Ser	GAG Glu 150	GGT Gly	GAA Glu	AGC Ser	ATC Ile	TGG Trp 155	GAG Glu	ATC Ile	ACT Thr	GAG Glu	AAA Lys 160	480
ATA Ile	CTA Leu	AAT Asn	TCG Ser	TTT Phe 165	GAG Glu	TAT Tyr	ACT Thr	TCG Ser 170	AGA Arg	TTT Phe	ACA Thr	AAA Lys	ACA Thr	AAA Lys 175	ACT Thr	528
TTA Leu	TAC Tyr	CAA Gln	TTC Phe 180	CTC Leu	TTC Phe	CTA Leu	GCT Ala	ACT Thr 185	TTC Phe	ATC Ile	AAT Asn	TGT Cys	GGA Gly 190	AGA Arg	TTC Phe	576
AGC Ser	GAT Asp	ATT Ile 195	AAG Lys	AAC Asn	GTT Val	GAT Asp	CCG Pro 200	AAA Lys	TCA Ser	TTT Phe	AAA Lys	TTA Leu 205	GTC Val	CAA Gln	AAT Asn	624
AAG Lys	TAT Tyr 210	CTG Leu	GGA Gly	GTA Val	ATA Ile	ATC Ile 215	CAG Gln	TGT Cys	TTA Leu	GTG Val	ACA Thr 220	GAG Glu	ACA Thr	AAG Lys	ACA Thr	672
AGC Ser 225	GTT Val	AGT Ser	AGG Arg	CAC His	ATA Ile 230	TAC Tyr	TTC Phe	TTT Phe	AGC Ser	GCA Ala 235	AGG Arg	GGT Gly	AGG Arg	ATC Ile	GAT Asp 240	720
CCA Pro	CTT Leu	GTA Val	TAT Tyr	TTG Leu 245	GAT Asp	GAA Glu	TTT Phe	TTG Leu 250	AGG Arg	AAT Asn	TCT Ser	GAA Glu	CCA Pro	GTC Val 255	CTA Leu	768
AAA Lys	CGA Arg	GTA Val 260	AAT Asn	AGG Arg	ACC Thr	GGC Gly	AAT Asn 265	TCT Ser 265	TCA Ser	AGC Ser	AAT Asn	AAA Lys 270	CAG Gln	GAA Glu	TAC Tyr	816
CAA Gln	TTA Leu 275	TTA Leu	AAA Lys	GAT Asp	AAC Asn	TTA Leu 280	GTC Val	AGA Arg	TCG Ser	TAC Tyr	AAT Asn	AAA Lys 285	GCT Ala	TTG Leu	AAG Lys	864

AAA AAT GCG CCT TAT TCA ATC TTT GCT ATA AAA AAT GGC CCA AAA TCT 912
 Lys Asn Ala Pro Tyr Ser Ile Phe Ala Ile Lys Asn Gly Pro Lys Ser
 290 295 300

CAC ATT GGA AGA CAT TTG ATG ACC TCA TTT CTT TCA ATG AAG GGC CTA 960
 His Ile Gly Arg His Leu Met Thr Ser Phe Leu Ser Met Lys Gly Leu
 305 310 315 320

ACG GAG TTG ACT AAT GTT GTG GGA AAT TGG AGC GAT AAG CGT GCT TCT 1008
 Thr Glu Leu Thr Asn Val Val Gly Asn Trp Ser Asp Lys Arg Ala Ser
 325 330 335

GCC GTG GCC AGG ACA ACG TAT ACT CAT CAG ATA ACA GCA ATA CCT GAT 1056
 Ala Val Ala Arg Thr Thr Tyr Thr His Gln Ile Thr Ala Ile Pro Asp
 340 345 350

CAC TAC TTC GCA CTA GTT TCT CGG TAC TAT GCA TAT GAT CCA ATA TCA 1104
 His Tyr Phe Ala Leu Val Ser Arg Tyr Tyr Ala Tyr Asp Pro Ile Ser
 355 360 365

AAG GAA ATG ATA GCA TTG AAG GAT GAG ACT AAT CCA ATT GAG GAG TGG 1152
 Lys Glu Met Ile Ala Leu Lys Asp Glu Thr Asn Pro Ile Glu Glu Trp
 370 375 380

CAG CAT ATA GAA CAG CTA AAG GGT AGT GCT GAA GGA AGC ATA CGA TAC 1200
 Gln His Ile Glu Gln Leu Lys Gly Ser Ala Glu Gly Ser Ile Arg Tyr
 385 390 395 400

CCC GCA TGG ATT GGG ATA ATA TCA CAG GAG GTA CTA GAC TAC CTT TCA 1248
 Pro Ala Trp Ile Gly Ile Ile Ser Gln Glu Val Leu Asp Tyr Leu Ser
 405 410 415

TCC TAC ATA AAT AGA CGC ATA TAAGTACGCA TTAAAGCATA AACACGCACT 1299
 Ser Tyr Ile Asn Arg Arg Ile 420

ATCCCGTTCT TCTCATGTAT ATATATATAC AGGCAACACG CAGATATAGG TGCGACGTGA 1359

ACAGTGAGCT GTATGTGCGC A 1380

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 423 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Pro Gln Phe Asp Ile Leu Cys Lys Thr Pro Pro Lys Val Leu Val
 1 5 10 15

Arg Gln Phe Val Glu Arg Phe Glu Arg Pro Ser Gly Glu Lys Ile Ala
 20 25 30

Leu Cys Ala Ala Glu Leu Thr Tyr Leu Cys Trp Met Ile Thr His Asn
 35 40 45

Gly Thr Ala Ile Lys Arg Ala Thr Phe Met Ser Tyr Asn Thr Ile Ile
 50 55 60

Ser Asn Ser Leu Ser Phe Asp Ile Val Asn Lys Ser Leu Gln Phe Lys
 65 70 75 80

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 34 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GAAGTTCCTA TTCTCTAGAA AGTATAGGAA CTTC

34

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 68 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GATCCCGGGC TACCATGGAG AAGTTCCTAT TCCGAAGTTC CTATTCTCTA GAAAGTATAG

60

GAACTTCA

68

"GATCCCGGGC TACCATGGAG AAGTTCCTAT TCCGAAGTTC CTATTCTCTA GAAAGTATAG GAACTTCA"